RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 09/997,807B
Source: 1FW16
Date Processed by STIC: 3/30/05

ENTERED

IFW16

RAW SEQUENCE LISTING DATE: 03/30/2005
PATENT APPLICATION: **US/09/997,807B** TIME: 15:21:33

Input Set : **D:\56446-20109.00 - SEQ Substitute (client).txt**
Output Set: **N:\CRF4\03302005\I997807B.raw**

<110> APPLICANT: Jay Short
Eric Mathur
William Michael Lafferty
Nelson Barton
Kevin Chow
<120> TITLE OF INVENTION: Method of Making a Protein Polymer and Uses of the Polymer
<130> FILE REFERENCE: 564462010900
<140> CURRENT APPLICATION NUMBER: 09/997,807B
<141> CURRENT FILING DATE: 2001-11-30
<150> PRIOR APPLICATION NUMBER: 60/250,426
<151> PRIOR FILING DATE: 2000-11-30
<160> NUMBER OF SEQ ID NOS: 37
<170> SOFTWARE: FastSEQ for Windows Version 4.0

p.6

<210> SEQ ID NO: 1
<211> LENGTH: 624
<212> TYPE: DNA
<213> ORGANISM: Pyrodictium abyssi
<400> SEQUENCE: 1

gtgaagtaca caaccctagc tatagcgggt attattgcct cggctgccgc cctcgccctc 60
ctagcaggct tcgccaccac ccagagcccc ctcaacagct tctacgccac cggtacagca 120
caggcagtaa gcgagccaat agacgtagaa agccacctcg gcagcataac ccccgcagcc 180
ggcgcacagg gcagtgacga cataggttac gcaatagtgt ggataaagga ccaggtcaat 240
gatgtaaagc tgaaggtgac cctgcgtaac gctgagcagc taaagcccta cttcaagtac 300
ctacagatac agataacaag cggctatgag acgaacagca cagctctagg caacttcagc 360
gagaccaagg ctgtgataag cctcgacaac cccagcgccg tgatagtact agacaaggag 420
gatatagcag tgctctatcc ggacaagacc ggttacacaa acacttcgat atgggtaccc 480
ggtgaacctg acaagataat tgtctacaac gagacaaagc cagtagctat actgaacttc 540
aaggccttct acgaggctaa ggagggtatg ctattcgaca gcctgccagt gatattcaac 600
ttccaggtgc tacaagtagg ctaa 624

<210> SEQ ID NO: 2
<211> LENGTH: 207
<212> TYPE: PRT
<213> ORGANISM: Pyrodictium abyssi
<400> SEQUENCE: 2

Val Lys Tyr Thr Thr Leu Ala Ile Ala Gly Ile Ile Ala Ser Ala Ala
1 5 10 15
Ala Leu Ala Leu Leu Ala Gly Phe Ala Thr Thr Gln Ser Pro Leu Asn
20 25 30
Ser Phe Tyr Ala Thr Gly Thr Ala Gln Ala Val Ser Glu Pro Ile Asp
35 40 45
Val Glu Ser His Leu Gly Ser Ile Thr Pro Ala Ala Gly Ala Gln Gly
50 55 60
Ser Asp Asp Ile Gly Tyr Ala Ile Val Trp Ile Lys Asp Gln Val Asn

RAW SEQUENCE LISTING DATE: 03/30/2005
PATENT APPLICATION: US/09/997,807B TIME: 15:21:33

Input Set : D:\56446-20109.00 - SEQ Substitute (client).txt
Output Set: N:\CRF4\03302005\I997807B.raw

```
57 65                  70                  75                  80
58 Asp Val Lys Leu Lys Val Thr Leu Arg Asn Ala Glu Gln Leu Lys Pro
59                  85                  90                  95
60 Tyr Phe Lys Tyr Leu Gln Ile Gln Ile Thr Ser Gly Tyr Glu Thr Asn
61             100                 105                 110
62 Ser Thr Ala Leu Gly Asn Phe Ser Glu Thr Lys Ala Val Ile Ser Leu
63         115                 120                 125
64 Asp Asn Pro Ser Ala Val Ile Val Leu Asp Lys Glu Asp Ile Ala Val
65     130                 135                 140
66 Leu Tyr Pro Asp Lys Thr Gly Tyr Thr Asn Thr Ser Ile Trp Val Pro
67 145                 150                 155                 160
68 Gly Glu Pro Asp Lys Ile Ile Val Tyr Asn Glu Thr Lys Pro Val Ala
69                  165                 170                 175
70 Ile Leu Asn Phe Lys Ala Phe Tyr Glu Ala Lys Glu Gly Met Leu Phe
71             180                 185                 190
72 Asp Ser Leu Pro Val Ile Phe Asn Phe Gln Val Leu Gln Val Gly
73         195                 200                 205
75 <210> SEQ ID NO: 3
76 <211> LENGTH: 513
77 <212> TYPE: DNA
78 <213> ORGANISM: Pyrodictium abyssi
80 <400> SEQUENCE: 3
81 gtgaagccta cggctctagc cctggctggt atcattgcct cggctgccga cctcgccctg      60
82 ctagcaggct tcgccaccac ccagagcccg ctcaacagct tctacgccac cggcacagca     120
83 gccgcaacaa gcgagccaat agacgtagag agccacctca gcagcatagc ccctgctgct     180
84 ggcgcacagg gcagccagga cataggctac ttcaacgtga ccgccaagga tcaagtgaac     240
85 gtgacaaaga taaaggtgac cctggctaac gctgagcagc taaagcccta cttcaagtac     300
86 ctacagatag tgctaaagag cgaggtagct gacgagatca aggccgtaat aagcatagac     360
87 aagcctagcg ccgtcataat actagacagc caggacttcg acagcaacaa cagagcaaag     420
88 ataagcgcca ctgcctacta cgaggctaag gagggcatgc tattcgacag cctaccgcta     480
89 atattcaaca tacaggtgct aagcgtcagc taa                                  513
91 <210> SEQ ID NO: 4
92 <211> LENGTH: 170
93 <212> TYPE: PRT
94 <213> ORGANISM: Pyrodictium abyssi
96 <400> SEQUENCE: 4
97 Val Lys Pro Thr Ala Leu Ala Leu Ala Gly Ile Ile Ala Ser Ala Ala
98  1               5                  10                  15
99 Asp Leu Ala Leu Leu Ala Gly Phe Ala Thr Thr Gln Ser Pro Leu Asn
100             20                  25                  30
101 Ser Phe Tyr Ala Thr Gly Thr Ala Ala Ala Thr Ser Glu Pro Ile Asp
102         35                  40                  45
103 Val Glu Ser His Leu Ser Ser Ile Ala Pro Ala Ala Gly Ala Gln Gly
104     50                  55                  60
105 Ser Gln Asp Ile Gly Tyr Phe Asn Val Thr Ala Lys Asp Gln Val Asn
106 65                  70                  75                  80
107 Val Thr Lys Ile Lys Val Thr Leu Ala Asn Ala Glu Gln Leu Lys Pro
108                 85                  90                  95
109 Tyr Phe Lys Tyr Leu Gln Ile Val Leu Lys Ser Glu Val Ala Asp Glu
```

RAW SEQUENCE LISTING DATE: 03/30/2005
PATENT APPLICATION: US/09/997,807B TIME: 15:21:33

Input Set : D:\56446-20109.00 - SEQ Substitute (client).txt
Output Set: N:\CRF4\03302005\I997807B.raw

```
110             100                 105                 110
111 Ile Lys Ala Val Ile Ser Ile Asp Lys Pro Ser Ala Val Ile Ile Leu
112         115                 120                 125
113 Asp Ser Gln Asp Phe Asp Ser Asn Asn Arg Ala Lys Ile Ser Ala Thr
114     130                 135                 140
115 Ala Tyr Tyr Glu Ala Lys Glu Gly Met Leu Phe Asp Ser Leu Pro Leu
116 145                 150                 155                 160
117 Ile Phe Asn Ile Gln Val Leu Ser Val Ser
118                 165                 170
120 <210> SEQ ID NO: 5
121 <211> LENGTH: 537
122 <212> TYPE: DNA
123 <213> ORGANISM: Pyrodictium abyssi
125 <400> SEQUENCE: 5
126 atgaggtaca cgaccctagc tctggccggc atagtggcct cggctgccgc cctcgccctg      60
127 ctagcaggct tcgccacgac ccagagcccg ctaagcagct tctacgccac cggcacagca     120
128 caagcagtaa gcgagccaat agacgtagag agccacctag acaacaccat agcccctgct     180
129 gccggtgcac agggctacaa ggacatgggc tacattaaga taactaacca gtcaaaagtt     240
130 aatgtaataa agctgaaggt gactctcgct aacgccgagc agctaaagcc ctacttcgac     300
131 tacctacagc tagtactcac aagcaacgcc actggcaccg acatggttaa ggctgtgcta     360
132 agcctcgaga agcctagcgc agtcataata ctagacaacg atgactacga tagcactaac     420
133 aagatacagc taaaggtaga agcctactat gaggctaagg agggcatgct attcgacagc     480
134 ctaccagtaa tactgaactt ccaggtactg agcgccgctt gcagtccctt gtggtga        537
136 <210> SEQ ID NO: 6
137 <211> LENGTH: 178
138 <212> TYPE: PRT
139 <213> ORGANISM: Pyrodictium abyssi
141 <400> SEQUENCE: 6
142 Met Arg Tyr Thr Thr Leu Ala Leu Ala Gly Ile Val Ala Ser Ala Ala
143  1               5                  10                  15
144 Ala Leu Ala Leu Leu Ala Gly Phe Ala Thr Thr Gln Ser Pro Leu Ser
145             20                  25                  30
146 Ser Phe Tyr Ala Thr Gly Thr Ala Gln Ala Val Ser Glu Pro Ile Asp
147         35                  40                  45
148 Val Glu Ser His Leu Asp Asn Thr Ile Ala Pro Ala Ala Gly Ala Gln
149     50                  55                  60
150 Gly Tyr Lys Asp Met Gly Tyr Ile Lys Ile Thr Asn Gln Ser Lys Val
151 65                  70                  75                  80
152 Asn Val Ile Lys Leu Lys Val Thr Leu Ala Asn Ala Glu Gln Leu Lys
153                 85                  90                  95
154 Pro Tyr Phe Asp Tyr Leu Gln Leu Val Leu Thr Ser Asn Ala Thr Gly
155             100                 105                 110
156 Thr Asp Met Val Lys Ala Val Leu Ser Leu Glu Lys Pro Ser Ala Val
157         115                 120                 125
158 Ile Ile Leu Asp Asn Asp Asp Tyr Asp Ser Thr Asn Lys Ile Gln Leu
159     130                 135                 140
160 Lys Val Glu Ala Tyr Tyr Glu Ala Lys Glu Gly Met Leu Phe Asp Ser
161 145                 150                 155                 160
162 Leu Pro Val Ile Leu Asn Phe Gln Val Leu Ser Ala Ala Cys Ser Pro
```

RAW SEQUENCE LISTING DATE: 03/30/2005
PATENT APPLICATION: US/09/997,807B TIME: 15:21:33

Input Set : D:\56446-20109.00 - SEQ Substitute (client).txt
Output Set: N:\CRF4\03302005\I997807B.raw

```
163                 165                 170                 175
164 Leu Trp
167 <210> SEQ ID NO: 7
168 <211> LENGTH: 395
169 <212> TYPE: DNA
170 <213> ORGANISM: Pyrodictium abyssi
172 <400> SEQUENCE: 7
173 agcttctacg ccaccggcac agcacaggca gtaagcgagc caatagacgt ggtaagcagc      60
174 ctcggtacgc taaatactgc cgctggtgca cagggtaagc agacgctagg agacataaca     120
175 atatatgcgc acaatgacgt gaacataaca aagctaaagg tcacgcttgc taacgctgca     180
176 cagctaagac catacttcaa gtacctgata ataaagctag taagcctgga cagcaacggc     240
177 aacgagtccg aggaaaaggg catgataact ctatggaagc cttacgccgt gataatacta     300
178 gaccatgaag atttcaacaa cgacatcgac aatgacggca acaatgacgc caagataagg     360
179 gttgtagcct actatgaggc taaggagggt atgct                               395
181 <210> SEQ ID NO: 8
182 <211> LENGTH: 131
183 <212> TYPE: PRT
184 <213> ORGANISM: Pyrodictium abyssi
186 <400> SEQUENCE: 8
187 Ser Phe Tyr Ala Thr Gly Thr Ala Gln Ala Val Ser Glu Pro Ile Asp
188  1               5                  10                  15
189 Val Val Ser Ser Leu Gly Thr Leu Asn Thr Ala Ala Gly Ala Gln Gly
190             20                  25                  30
191 Lys Gln Thr Leu Gly Asp Ile Thr Ile Tyr Ala His Asn Asp Val Asn
192         35                  40                  45
193 Ile Thr Lys Leu Lys Val Thr Leu Ala Asn Ala Ala Gln Leu Arg Pro
194     50                  55                  60
195 Tyr Phe Lys Tyr Leu Ile Ile Lys Leu Val Ser Leu Asp Ser Asn Gly
196 65                  70                  75                  80
197 Asn Glu Ser Glu Glu Lys Gly Met Ile Thr Leu Trp Lys Pro Tyr Ala
198                 85                  90                  95
199 Val Ile Ile Leu Asp His Glu Asp Phe Asn Asn Asp Ile Asp Asn Asp
200             100                 105                 110
201 Gly Asn Asn Asp Ala Lys Ile Arg Val Val Ala Tyr Tyr Glu Ala Lys
202         115                 120                 125
203 Glu Gly Met
204     130
206 <210> SEQ ID NO: 9
207 <211> LENGTH: 372
208 <212> TYPE: DNA
209 <213> ORGANISM: Pyrodictium abyssi
211 <400> SEQUENCE: 9
212 agcttctacg ccaccggcac agcagaggca acaagcgagc caatagacgt tgtaagcaac      60
213 cttaacacgg ccatagcccc tgctgccggc gcccagggca gcgtgggcat aggcagcata     120
214 acaatagaga acaagactga cgtgaacgtt gtgaagctga agataaccct cgccaacgct     180
215 gagcagctaa agccctactt cgactaccta cagatagtgc taaagagcgt tgacagcaac     240
216 gagatcaagg ctgtgctaag cctcgagaag cccagcgcag tcataatact ggacaacgag     300
217 gacttccagg gcggcgacaa ccagtgccag atagacgcca ccgcctacta cgaggctaag     360
218 gagggtatgc ta                                                       372
```

RAW SEQUENCE LISTING DATE: 03/30/2005
PATENT APPLICATION: US/09/997,807B TIME: 15:21:33

Input Set : D:\56446-20109.00 - SEQ Substitute (client).txt
Output Set: N:\CRF4\03302005\I997807B.raw

<210> SEQ ID NO: 10
<211> LENGTH: 124
<212> TYPE: PRT
<213> ORGANISM: Pyrodictium abyssi
<400> SEQUENCE: 10

```
Ser Phe Tyr Ala Thr Gly Thr Ala Glu Ala Thr Ser Glu Pro Ile Asp
 1               5                  10                  15
Val Val Ser Asn Leu Asn Thr Ala Ile Ala Pro Ala Ala Gly Ala Gln
            20                  25                  30
Gly Ser Val Gly Ile Gly Ser Ile Thr Ile Glu Asn Lys Thr Asp Val
        35                  40                  45
Asn Val Val Lys Leu Lys Ile Thr Leu Ala Asn Ala Glu Gln Leu Lys
    50                  55                  60
Pro Tyr Phe Asp Tyr Leu Gln Ile Val Leu Lys Ser Val Asp Ser Asn
65                  70                  75                  80
Glu Ile Lys Ala Val Leu Ser Leu Glu Lys Pro Ser Ala Val Ile Ile
                85                  90                  95
Leu Asp Asn Glu Asp Phe Gln Gly Gly Asp Asn Gln Cys Gln Ile Asp
            100                 105                 110
Ala Thr Ala Tyr Tyr Glu Ala Lys Glu Gly Met Leu
        115                 120
```

<210> SEQ ID NO: 11
<211> LENGTH: 448
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: consensus sequence
<400> SEQUENCE: 11

```
tgagacccta gctgcggatt gcctcggctg ccgcctcgcc ctctagcagg cttcgccaca      60
cccagagccc ctacagcttc tacgccaccg gcacagcaca ggcagtaagc gagccaatag     120
acgtagaaag ccacctcaca catagcccct gctgccggcg cacagggcag caggacatag     180
gctacataaa ataacaagat agtgaacgta taaagctgaa ggtgaccctg ctaacgctga     240
gcagctaaag ccctacttca agtacctaca gatagtgcta aaagcgacag caggcacacg     300
agaaggcgtg ataagcctcg agaagcctag cgccgtcata atactagaca acgaggactt     360
cgaagcacaa cagaaagaga agcaatagcc tactacgagg ctaaggaggg tatgctattc     420
gacagcctcc tatataactc aggtctgt                                       448
```

<210> SEQ ID NO: 12
<211> LENGTH: 140
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: consensus sequence
<400> SEQUENCE: 12

```
Val Lys Thr Leu Ala Leu Ala Gly Ile Ile Ala Ser Ala Ala Leu Ala
 1               5                  10                  15
Leu Leu Ala Gly Phe Ala Thr Thr Gln Ser Pro Leu Ser Phe Tyr Ala
            20                  25                  30
Thr Gly Thr Ala Gln Ala Val Ser Glu Pro Ile Asp Val Glu Ser His
        35                  40                  45
```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 03/30/2005
PATENT APPLICATION: **US/09/997,807B** TIME: 15:21:34

Input Set : **D:\56446-20109.00 - SEQ Substitute (client).txt**
Output Set: **N:\CRF4\03302005\I997807B.raw**

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:20; N Pos. 6,7,8,9,10

VERIFICATION SUMMARY DATE: 03/30/2005
PATENT APPLICATION: **US/09/997,807B** TIME: 15:21:34

Input Set : **D:\56446-20109.00 - SEQ Substitute (client).txt**
Output Set: **N:\CRF4\03302005\I997807B.raw**

L:375 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:379 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:20
L:380 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20 after pos.:0